THE DURBIN-WATSON TEST

Suppose we have a time series regression model relating a "dependent" time series $\{y_t\}$ to the "independent" time series $\{x_{1t}\}, \ldots, \{x_{pt}\}$. The model is

$$y_t = \beta_o + \beta_1 x_{1t} + \cdots + \beta_p x_{pt} + \varepsilon_t \quad , \quad t = 1, 2, \ldots, n \quad ,$$

where $\{\varepsilon_t\}$ is a time series of "errors", or "disturbances". Such models are useful for both explanatory and forecasting purposes. The parameters $\beta_0, \beta_1, \ldots, \beta_p$ may be estimated by least-squares. In practice, it often happens that the errors are not independent (as assumed in standard regression models) but instead are *autocorrelated*. Such error autocorrelation, or "serial correlation", has many underirable but correctable consequences (e.g., the least-squares estimates sub-optimal, standard confidence intervals for β are incorrect, the error term is forecastable). Thus, it is highly desirable to try to detect error autocorrelations.

The **Durbin-Watson Test** for serial correlation assumes that the ε_t are stationary and normally distributed with mean zero. It tests the null hypothesis H_o that the errors are uncorrelated against the alternative hypothesis H_1 that the errors are $AR(1)$. Thus, if ρ_s are the error autocorrelations, then we have $H_o : \rho_s = 0$ $(s > 0)$, and $H_1 : \rho_s = \rho^s$ for some nonzero ρ with $|\rho| < 1$. To test H_o against H_1, get the least squares estimates $\hat{\beta}$ and residuals $e_1, \ldots, e_n$. The test statistic is

$$d = \sum_{t=2}^{n} (e_t - e_{t-1})^2 / \sum_{t=1}^{n} e_t^2 \quad .$$

Note that ignoring "end effects", we have $d \approx 2(1 - r_1)$, where r_1 is the sample ACF of the residuals at lag 1. If the errors are white noise, d will be close to 2. If the errors are strongly autocorrelated, d will be far from 2. The exact procedure for deciding whether a given value of d is significant is somewhat complicated, and is described, for example, in Draper and Smith, Page 163. In some cases, the test can be "inconclusive," i.e., H_o is neither accepted nor rejected.

Since its development in 1951, the test has been found to be extremely useful, especially for the analysis of economic time series. It does, however, suffer from a number of shortcomings, some of which are as follows. First, the form of the model (i.e., the dimension p and the explanatory variables $x_{1t}, \ldots, x_{pt}$) is assumed known. In practice, this is rarely the case, and instead a data based procedure must be employed to "identify" the model. Second, the test is sometimes inconclusive, as mentioned above. Third, the $AR(1)$ alternative hypothesis is by no means the only way in which the null hypothesis may fail. Suppose, for example, the errors are in fact $MA(1)$, or perhaps even some nonstationary series such as a random walk. The Durbin-Watson test can have very low power against such alternatives (i.e., it can fail to detect them).